

## CLAIMS

1. A method for ranking a set of documents, comprising the steps of:
  - gathering context information from the documents;
  - 5 - generating at least one rank criterion from the context information; and
  - ranking the documents, based on the at least one rank criterion.
2. The method according to Claim 1, further comprising re-ranking an existing ranked result set of documents.
- 10 3. The method according to Claim 2, wherein said step of gathering context information comprises extracting lexical affinities from the documents.
4. The method according to Claim 2, wherein said step of gathering context  
15 information comprises extracting features from the documents.
5. The method according to Claim 2, wherein said step of gathering context information comprises extracting word frequency statistics from the documents.
- 20 6. The method according to any of Claims 1 to 5, further comprising the step of weighting of the context information by a weighting function.
7. The method according to Claim 6, further comprising the step of utilizing discrete ranking levels in said weighting step.
- 25 8. A method for re-ranking an existing set of text documents, comprising the steps of:
  - detecting lexical affinity terms contained in the documents;
  - presenting the lexical affinity terms to a user;
  - gathering user preferences for the lexical affinity terms; and
  - 30 - re-ranking the documents based on the user preferences.

9. A method for re-ranking an existing set of text documents, comprising the steps of:
- detecting feature terms contained in the documents;
  - presenting the feature terms to a user;
  - gathering user preferences for the feature terms; and
  - 5 - re-ranking the documents based on the user preferences.
10. A method for re-ranking an existing set of text documents, comprising the steps of:
- creating word frequency statistics from the documents;
  - presenting the words with a minimum frequency to a user;
  - 10 - gathering user preferences for the presented words of a minimum frequency; and
  - re-ranking the documents based on the user preferences.
11. A method according to any of Claims 8 to 10, wherein the re-ranking is based on the original ranking position of the documents.
- 15 12. A method according to Claim 1, wherein said step of ranking the documents comprises using the following ranking and weighted ranking equations or their equivalence:
- ranking equation -
- 20  $fd(x_1, \dots, x_n) = R_d$  if  $x_1, \dots, x_n$  are elements of  $T_d$ , and
- $fd(x_1, \dots, x_n) = 0$  if  $x_1, \dots, x_n$  are not elements of  $T_d$ ,
- wherein  $R_d$  is an “absolute” rank value of a given document “d” that has resulted from a search, and  $T_d = (x_1, \dots, x_n)$  is a tuple of context terms that are contained in the document “d”;
- 25 weighted ranking equation -
- $[2a f(x_1, \dots, x_a) + (a+b) f(x_1, \dots, x_{a+b}) + (a+b+c) f(x_1, \dots, x_{a+b+c})] / (4a+2b+c)$
- wherein it calculates the relevance of a document with respect to the context terms  $x_1, \dots, x_m$  when  $a, b$  and  $c$  are the number of terms that have been assigned a high ( $a$ ), medium ( $b$ ) and low ( $c$ ) relevance and  $f(x_1, \dots, x_a)$ ,
- 30  $f(x_1, \dots, x_{a+b})$  and  $f(x_1, \dots, x_{a+b+c})$  are partial relevance functions of the document with respect to a subset of the context terms.

13. A system for ranking a set of documents, comprising:
- means for gathering context information from the documents;
  - means for generating at least one rank criterion from the context information; and
  - means for ranking the documents, based on the at least one rank criterion.
14. A system according to Claim 13, further comprising means for re-ranking an existing ranked result set of documents.
15. A system according to Claim 13, further comprising means for extracting lexical affinities from the documents in order to obtain the context information.
16. A system according to Claim 16, further comprising means for weighting of the context information by a weighting function.
17. A computer-readable program storage medium which stores a program for executing a method for ranking a set of documents, the method comprising the steps of:
- gathering context information from the documents;
  - generating at least one rank criterion from the context information; and
  - ranking the documents, based on the at least one rank criterion.
18. The computer-readable program storage medium according to Claim 17, further comprising re-ranking an existing ranked result set of documents.
19. The computer-readable program storage medium according to Claim 18, wherein said step of gathering context information comprises extracting lexical affinities from the documents.

20. The computer-readable program storage medium according to Claim 18, wherein said step of gathering context information comprises extracting features from the documents.
- 5 21. The computer-readable program storage medium according to Claim 18, wherein said step of gathering context information comprises extracting word frequency statistics from the documents.
- 10 22. The computer-readable program storage medium according to any of Claims 17 to 21, further comprising the step of weighting of the context information by a weighting function.
- 15 23. The computer-readable program storage medium according to Claim 22, further comprising the step of utilizing discrete ranking levels in said weighting step.
- 20 24. A computer-readable program storage medium which stores a program for executing a method for re-ranking an existing set of text documents, comprising the steps of:
- detecting lexical affinity terms contained in the documents;
  - presenting the lexical affinity terms to a user;
  - gathering user preferences for the lexical affinity terms; and
  - re-ranking the documents based on the user preferences.
- 25 25. A computer-readable program storage medium which stores a program for executing a method for re-ranking an existing set of text documents, comprising the steps of:
- detecting feature terms contained in the documents;
  - presenting the feature terms to a user;
  - gathering user preferences for the feature terms; and
  - 30 - re-ranking the documents based on the user preferences.

26. A computer-readable program storage medium which stores a program for executing a method for re-ranking an existing set of text documents, comprising the steps of:

- creating word frequency statistics from the documents;
- presenting the words with a minimum frequency to a user;
- gathering user preferences for the presented words of a minimum frequency; and
- re-ranking the documents based on the user preferences.

27. The computer-readable program storage medium according to any of Claims 24 to 26, wherein the re-ranking is based on the original ranking position of the documents.

28. The computer-readable program storage medium according to Claim 17, wherein said step of ranking the documents comprises using the following ranking and weighted ranking equations or their equivalence:

ranking equation -

$fd(x_1, \dots, x_n) = R_d$  if  $x_1, \dots, x_n$  are elements of  $T_d$ , and

$fd(x_1, \dots, x_n) = 0$  if  $x_1, \dots, x_n$  are not elements of  $T_d$ ,

wherein  $R_d$  is an "absolute" rank value of a given document "d" that has resulted from a search, and  $T_d = (x_1, \dots, x_n)$  is a tuple of context terms that are contained in the document "d";

weighted ranking equation -

$[2a f(x_1, \dots, x_a) + (a+b) f(x_1, \dots, x_{a+b}) + (a+b+c) f(x_1, \dots, x_{a+b+c})] / (4a+2b+c)$

wherein it calculates the relevance of a document with respect to the context terms  $x_1, \dots, x_m$  when  $a, b$  and  $c$  are the number of terms that have been assigned a high ( $a$ ), medium ( $b$ ) and low ( $c$ ) relevance and  $f(x_1, \dots, x_a)$ ,  $f(x_1, \dots, x_{a+b})$  and  $f(x_1, \dots, x_{a+b+c})$  are partial relevance functions of the document with respect to a subset of the context terms.